Project Goals

wkolcan2, guoyang2, bwl3, skavi2

For our CS 225 final project, we will be using the open-source dataset OpenFlights. Specifically for this project, we are interested in the airports and flight route data. We chose this dataset because we found it to be very interesting as well as practical. The two algorithms that we will implement are Dijkstra's algorithm and the PageRank algorithm because we feel that they mesh well with the OpenFlights dataset.

We believe Pagerank can provide insights into our target dataset. Typically, Pagerank uses incoming and outgoing links to various web pages to rank the relative importance of each web page. By replacing web pages with airports, and links with flights, we can use the same algorithm to rank the relative importance of airports. We will output a list giving each airport an importance value (all of which will sum to 1).

Dijkstra's algorithm is a natural choice for this data as well. The algorithm finds the shortest path between two nodes in linear time. When each node is an airport, and each edge a route, this has the very practical application of finding the shortest flight path between two locations. We will use this algorithm to find the shortest routes between two inputted airports and provide the travel time and/or distance.

For the traversal method of our graph, we have decided to go with breadth-first search. This is because it is more suitable for how PageRank is implemented. Which makes sense since we will be focusing on the number of flights entering and leaving an airport. Dijkstra's algorithm is also relatively easy to implement with a breadth-first search method.

In conclusion, we will be using the OpenFlights dataset. With this dataset, we will implement both the PageRank and Dijkstra's algorithms. Finally, we will be using a breadth-first search to traverse through the graph that we make.

(Photograph of Group Members)